Volume 12 Number 3 1984 Nucleic Acids Research

Intracellular forms of viral DNA consistent with a model of reverse transcriptional replication of the cauliflower mosaic virus genome

Yves Marco and Stephen H.Howell*

Biology Department CO16, University of California, San Diego, La Jolla, CA 92093, USA

Received 27 July 1983; Revised and Accepted 15 December 1983

ABSTRACT

Five unencapsidated, intracellular forms of cauliflower mosaic virus DNA which accumulate in infected turnip leaves are described. The forms are double stranded (or partially double stranded), and one (form iv) is a covalently-closed circular form of the full-length genome. Form i is an open form composed of full-length genome strands (7.6 kilobases) similar to encapsidated viral DNA. Form iii appears as a diffuse band on gels and is composed of two approximately half length strands (~3.8 kb) spanning the region of the genome between the two site-specific DNA breaks (Δ1 and Δ2). Form ii is composed of a full length strand and an approximately half-length strand as described for form iii DNA. Form v is a small form (~0.7 kilobases in undenatured form) that maps adjacent to the α-strand break (Δ1) and may be a fold-back form. These forms appear to be intermediates in cauliflower mosaic virus DNA replication and the properties of these forms are consistent with possible intermediates in a model of reverse transcriptional replication of the viral genome.

INTRODUCTION

Cauliflower mosaic virus (CaMV) is a plant virus with a circular, double-stranded DNA genome of about 8 kb. A peculiar feature of the genome is the presence of site-specific, single-strand breaks in the DNA extracted from virus particles. All viral isolates have a single break or discontinuity in the α-strand (Δ1) which has been set as the zero point in the map of the CaMV genome (1) and one or more breaks, depending on the isolate, in the opposite or β-strand (2-4). The termini of the DNA strands at the breaks are not abutted, but the ends overlap each other to different extents (5). The 5'-ends at each break are unique (6), but the 3'-ends terminate at any one of the several closely-spaced sites (5). CaMV DNA is also found in a covalently closed, unencapsidated form in the nuclei of infected cells (7,8). This form, termed a "minichromosome", possesses nucleosome structure and is transcriptionally active (7).

The CaMV genome is asymmetrically transcribed from α-strand (9) to yield several RNAs including a large 35S transcript which is slightly longer than

the full length viral genome. Transcription of the 35S RNA overlaps itself at its ends yielding terminal direct repeats in the primary transcript of about 180b (10,11). Hull and Covey (12) and Guilley et al. (13) have recently proposed that the large RNA transcript is a template for reverse transcriptional replication of CaMV DNA. They suggested that first strand CaMV DNA synthesis like retrovirus replication, is primed near the 5'-terminus of the template RNA by a tRNA that binds to the template at a site adjacent to the α-strand break (Δ1) in the corresponding DNA strand.

In this paper, we describe several unencapsidated, intracellular forms of the CaMV genome which may be intermediates (or are derived from intermediates) in viral DNA replication. The presumed intermediates appear to accumulate because of the postulated discontinuous process of CaMV DNA synthesis in which DNA strands which are initiated at the break sites in the CaMV genome, but are terminated at interruptions in the template -- in the RNA template in the synthesis of the first DNA strand and in the first strand DNA template in the synthesis of the second DNA strand. Covey et al. (14) and Hull and Covey (15) have recently described a number of different CaMV DNA fragments which accumulate in leaves of infected plants. Some of the forms which they have described are consistent with the notion that they are replication intermediates. Others appear to breakdown products which arise from double strand breaks in genome at the site of the single strand breaks (15). Since the complexity of the intracellular DNA fragment pattern appears to relate to the number of single strand breaks in the genome, we have chosen to study a viral isolate CM4-184 which has only two such breaks (3), one in either DNA strand.

EXPERIMENTAL PROCEDURES

PLANTS AND CLONED VIRAL DNA

Turnips (Brassica rapa L.C.V. Just Right) were grown under greenhouse conditions and usually harvested 2-3 weeks post-inoculation. Plasmid pLW414 containing the complete CaMV genome (of the CM4-184 viral isolate) inserted into the SalI site of pBR322 has been described (16).

DNA Extraction of CaMV Infected Leaves

DNA was extracted with slight modifications of the method described by Mennisier et al. (8). After RNase and Proteinase K treatment and Et OH precipitation DNA was usually banded on a CsCl gradient (ref. index = 1.400), and before analysis, was either precipitated with 3 volumes of 70% EtOH and washed twice with 70% EtOH or was dialysed against TE buffer (10mM Tris, 1mM EDTA, pH 8.0).

Hybridization Analysis

DNA was subjected to electrophoresis under standard conditions in non-denaturing agarose gels containing Tris-acetate buffer or under fully denaturing conditions in methylmercuric- hydroxide containing agarose gels according to Bailey and Davidson (17). Gel separated DNAs were transferred onto nitrocellulose paper according to the method of Southern (18) and hybridized by the procedure of Wahl et al. (19).

Labeling and Sequencing of DNA Fragments

Restriction endonucleases were purchased and used according to the recommendations of the manufacturer. DNA restriction fragments were labeled by nick translation following a procedure modified by Rigby et al. (20).

RESULTS

DNA from infected plant leaves appears in a variety of forms when analyzed by electrophoresis on agarose gels (Fig. 1). These DNAs apparently are all unencapsidated, forms, because Menissier et al. (8) have shown that the DNA isolation procedure, which we have used, does not extract DNA from intact virus particles. A variety of slow migrating DNA forms, labeled i-iv, and a rapidly migrating component, indicated as form v are observed in native or undenatured DNA samples (Fig. 1, Nat lane). Upon heat denaturation of total DNA, three major bands, labeled DNA A-C (Fig. 1, Den lane) and several minor bands are resolved. DNA A comigrates on methylmercurichydroxide gels (not shown) with full length, single strand CaMV DNA (~7.6 kilobases) derived by excision of the entire CaMV genome from pLW414 by cleavage with Sal I (Fig. 4). DNA B migrates here and on fully denaturing gels as a single strand DNA of 3.8 kb.

DNA C has unusual migration properties on various gel systems. In experiments where we have isolated form v from sucrose density gradient fractions, we have found the DNA C is derived from native form v. Nonetheless, upon denaturation, DNA C migrates slightly more rapidly on non-denaturing gels than form v (Fig. 1, Den lane). Form v migrates equivalent to a double strand linear DNA fragment of ~710 basepairs and upon denaturation migrates equivalent to a double strand form of 670 basepairs. On fully denaturing methylmercurichydroxide containing gels (not shown), DNA C migrates at nearly twice its expected size (1.3 kilobases). This suggests that a principal conformation of form v is a fold-back molecule that snaps back on itself upon renaturation. DNA D also appears to be derived from form v; however it is usually a minor form and we have not studied it extensively.

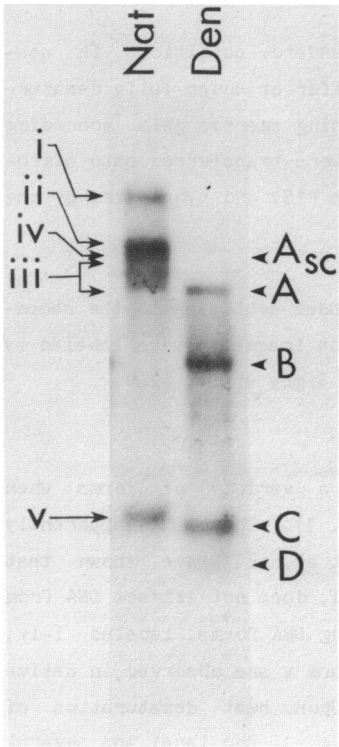

Figure 1. DNAs from CaMV leaves separated on a non-denaturing 1.5% agarose gel. DNA (10 µg) in native (Nat) or heat denatured (Den) samples was transferred to nitrocellulose paper and hybridized to nick-translation labeled CaMV DNA, BglII-3 and -4 fragments from pLW414 (see Fig. 4). Designations of DNA bands are discussed in the text.

The nature of DNA A_{sc} (Fig. 1, Den lane) and form iv (Fig. 1, Nat lane) was revealed when CsCl-ethidium bromide density gradient centrifugation analysis was carried out on infected leaf DNA. Viral DNAs from fractions spanning the region of the gradient from the banding densities of supercoiled (SC) to open circular (OC) and linear (L) DNA were examined (Fig. 2). DNA A_{sc} bands with SC DNA and all other forms band with OC or L DNA. DNA from the SC fraction comigrates on nondenaturing gels with form iv DNA as does the heat denatured DNA A_{sc} (Fig. 1, Nat lane). This is consistent with the expected snap-back behavior of a covalently closed circular DNA molecule. We assume that form iv DNA is full genome length, consistent with the observations of others (7,8). DNAs C and D do not sharply band (Fig. 2), probably due to their small size. Nonetheless, DNA C does not seem to be disposed toward SC fractions, indicating that it is probably not a supercoiled form.

To examine more carefully the slow migrating DNA forms observed in Fig. 1 (Nat lane) and to determine from which forms DNAs A-C were derived, we subjected infected leaf DNA to two dimensional gel analysis. To do so, native

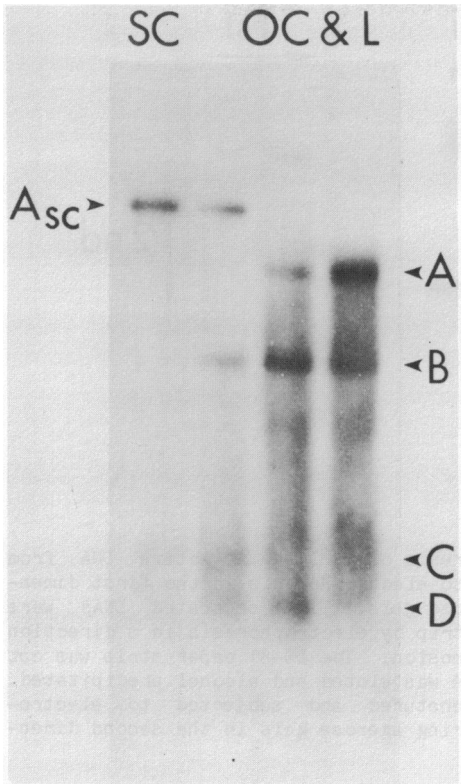

Figure 2. Analysis of viral DNA forms from an ethidium bromide-CsCl density gradient. DNA from four fractions spanning the region of the gradient from covalently closed or supercoiled (SC) DNA to open circular (OC) and linear (L) forms was heat denaturated and separated on non-denaturing 1.5% agarose gels. The DNA was transferred to nitrocellulose filters and hybridized to nick-translated CaMV DNA, BglII-3 and -4 fragments.

DNAs were separated on 1% non-denaturing agarose gels in the first dimension, and the DNAs from fractions of the gel in the first dimension were denatured and separated on 1.5% non-denaturing agarose gels in the second dimension. Forms i and ii appeared as discrete bands in the first dimension, but form iii was more diffuse (Fig. 3). Form i was mostly composed of DNA A. Form ii was a fairly equal mix of DNA A and B, and the region containing form iii was composed almost exclusively of DNA B (in the center fraction). Particularly, in the region containing form iii, many minor or less abundant DNA fragments were also found.

Since DNAs B and C are less than full genome length, we mapped the regions to which these DNAs correspond on the CaMV genome. The mapping was carried out by hybridizing various ^{32}P-labeled restriction fragments from cloned CaMV DNA, pLW414, to blots of gel-separated DNAs (Fig. 4). The blots are arranged so that the order of probes used moves counter-clockwise around the CaMV genome starting from the d-strand discontinuity site (Δ1). Using the ClaI pLW414 DNA probes, it can be observed that DNA B hybridizes to ClaI-1 and

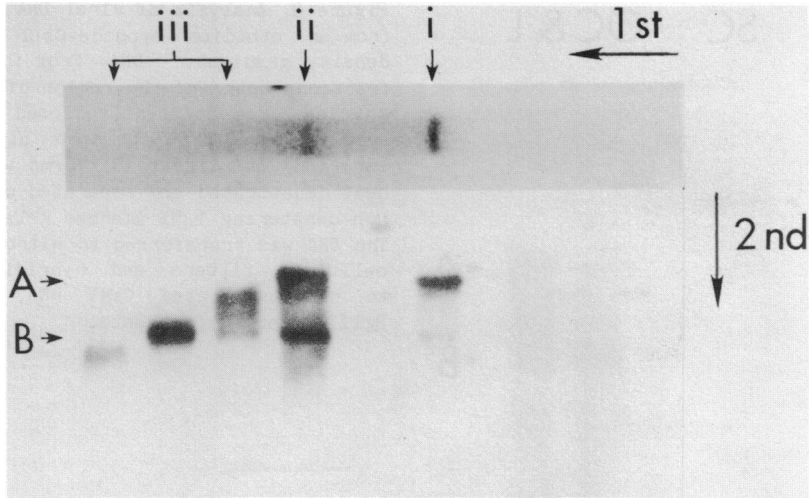

Figure 3. Two dimensional gel electrophoresis of CaMV DNA. Native DNA from infected leaves was subjected to electrophoresis in the first dimension on 1% non-denaturing agarose gels. The separated DNAs were collected on a DE-81 paper strip by electrophoresis in a direction perpendicular to the first dimension. The DE-81 paper strip was cut into 1 cm pieces, and the DNA was eluted and alcohol precipitated. The eluted samples were heat denatured and subjected to electrophoresis on 1.5% non-denaturing agarose gels in the second dimension.

-2. (DNA C is not readily seen in the DNA preparation used in the ClaI blots.) With the BglII probes, DNA C hybridizes to BglII-5 and -3. DNA B was not as clearly resolved in the BglII hybridization experiments, but it can be seen that DNA B hybridizes with the small BglII-5 fragment and not with BglII-6 and -4, which defines a border for DNA B near or at the /\1 site. DNA A hybridizes to all probes, expected for a full genome length DNA. Combining the fragment sizing data with the hybridization results and assuming that DNAs B and C terminate at /\1, DNA B extends to /\2, a distance of ~3810 basepairs (1), and DNA C extends approximately to the cap site of the 35S RNA. The cap site of 35S RNA is ~600 basepairs from 5' end of the (-strand at /\1 in the Cabb B-S viral isolate (10) and the length of DNA C on non-denaturing gels is ~670 basepairs.

When strand separated probes (CaMV DNA fragments cloned into M13mp8) were used to determine the strandedness of these DNA fragments, it was found that the DNAs A-C hybridize to probes from both strands (data not shown). Hence, we conclude that these DNAs might be double stranded or partially double stranded forms.

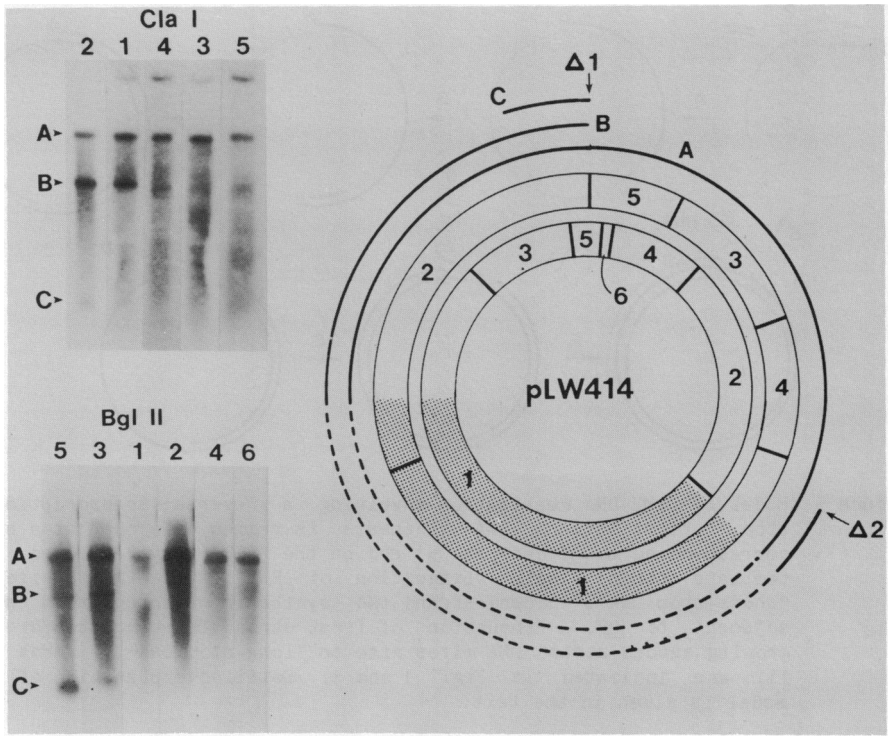

Figure 4. Mapping of DNAs A-C on the CaMV genome. The entire CaMV genome (CM4-184 isolate) is inserted in the SalI site of pBR322 (indicated by shaded sector) in pLW414. Inner ring on pLW414 map shows position of BglII sites and outer ring indicates ClaI sites. Sites corresponding to discontinuities in the α- and β-strand strands of virion DNA are indicated as Δ1 and Δ2, respectively. Map positions of DNAs A-C were deduced from the hybridization of ^{32}P-labeled ClaI fragments 1-5 (upper left panel) and BglII fragments 1-6 (lower left panel) of pLW414 to heat denatured infected leaf DNAs separated on 1.5% agarose gels.

DISCUSSION

We have described several unencapsidated, intracellular forms of CaMV DNA which accumulate in leaves of infected plants. We suggest that these forms are intermediates (or derived from intermediates) in the process of CaMV DNA replication. If CaMV DNA synthesis was a continuous process one might not expect to find the accumulation of discrete replication intermediates. However recently a reverse transcriptional model of CaMV DNA synthesis involving a number of discontinuous steps has been formulated by Hull and Covey (12) and by Guilley *et al.* (13) and discussed by Howell *et al.* (21). A version of

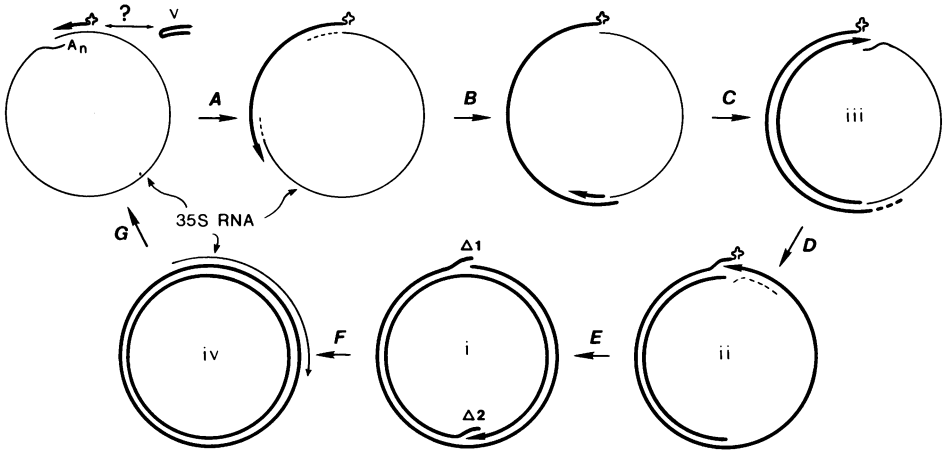

Figure 5. Model of CaMV DNA replication involving a reverse transcriptional step. First strand DNA synthesis is primed by a tRNA at a site corresponding to $\Delta 1$ near the 5' end on the 35S polyadenylated RNA template (thin line). Exhaustion of RNA template gives rise to "short-stop DNA". Second strand DNA synthesis starts at a site adjacent to $\Delta 2$. Exhaustion of first strand DNA template for the growing second strand DNA gives rise to "long-stop DNA". Forms i-iii are indicated in Fig. 1 and 3. Detailed explanation of the model is given in the text.

that model shown in Fig. 5 accommodates all the DNA forms we have described in this study.

An essential feature of that model is the role of the single strand breaks, found in virion DNA, as sites for the initiation of α and β-strand DNA synthesis. To simplify our understanding of the replication process we have chosen to study DNA from a viral isolate, CM4-184, containing only two breaks, $\Delta 1$ in the α-strand and $\Delta 2$ in the β-strand (3). In this model, first-or α-strand synthesis is primed by a tRNA at a site on the template 35S RNA corresponding to $\Delta 1$. A region of the CaMV genome complementary to 15 of the first 16 nucleotides at the 3'-end of wheat germ methionine tRNA and oriented in the appropriate polarity for priming has been located just adjacent to $\Delta 1$ (13). Because first strand DNA synthesis is initiated near the 5' end of the 35S RNA (~600 bases from the cap site, 10), the growing DNA strand quickly exhausts its template. Covey et al. (23) have recently described a small, composite molecule containing RNA and DNA that fulfills almost all of the expected features of the extended primer. The molecule called sa-DNA is 725

bases long, maps to the appropriate region on the genome as indicated in Fig. 5, is complementary to 35 S RNA, appears to be noncovalently attached to polyadenylated RNA from infected leaves (at least, it copurifies with poly[A] RNA) and is covalently linked to about 100 bases of RNA (presumably the tRNA primer).

As shown in Fig. 5, step A, we suggest that either one of two fates may befall the extended primer. The growing end may successfully transfer from 5'-end of the 35S RNA to the 3' end, permitting the further elongation of the first DNA strand (Fig. 5, step A). Such a transfer is possible by virtue of the fact that the 180 bases at the 5' end of 35S RNA are also found at the 3' end of the transcribed portion of the 35S RNA, excluding the poly[A] tail (10,11). Alternatively, the DNA may fold back and initiate second strand DNA synthesis on itself, as is suggested for the structure of form v. Covey et al. (23) also find in the DNA fractions of infected plants a double-stranded DNA structure similar to our form v. However, upon denaturation, their DNA falls apart in three species of 625, 575 and 200 bases (sb-, sc- and sd-DNA). We cannot exclude the possibility that some portion of our form V molecules might have single strand cuts and be composed of smaller single strand DNA pieces as described by Covey et al. (14). Such molecules may be the source of the minor DNA fragment D and other, as yet, undetected DNA fragments in our preparations. However, most of our form v molecules behave as single molecular species following denaturation as shown in Fig. 1.

Elongation of the first DNA strand presumably would be accompanied by destruction of the RNA template in regions of RNA-DNA hybrid by an RNase-H-like activity (Fig. 5, step A). Second strand DNA synthesis would commence just as first strand synthesis proceeds beyond /\2, the break site in the +B-strand (Fig. 5, step B). Guilley et al. (13) have suggested that second strand synthesis might be initiated, as is thought to be the case in retroviruses (24) at a purine-rich tract which in the CaMV genome is just upstream from /\2 (1). Second strand synthesis would continue until the elongating chain exhausted its first strand template, reaching the /\1 site (Fig. 5, step C). The truncated second strand appears to extend from /\2 to /\1 and is found in forms ii and iii described in this paper. Unfortunately, we cannot accurately map or sequence this second strand DNA to determine precisely its limits because although we have identified forms ii and iii, we have not isolated them. Form iii does not migrate on gels as a discrete band (Fig. 3) presumably because of variability in the length of first strand DNA. Further elongation of the first strand (Fig. 5, step D) produces form ii composed of a

complete first strand and the truncated second strand. Finally, through transfer of the growing end of the second DNA strand from the 5' end of the first strand template to its 3' end, the second DNA strand can complete its replication giving rise to form i DNA (Fig. 5, step E). Transfer of strands to elongate the second strand is made possible, as with the first strand, by overlapping sequences at the ends of the (-)-strand which may serve to bridge the two ends of the template DNAs. Here the overlapping sequences occur at $\Delta 1$ which has a 8-9 base overlap in the DNA of mature virions (5), and the overlap in the replicating DNA may be longer if the tRNA primer remains linked to the 5' end of the DNA at $\Delta 1$.

The origin of form iv DNA, the covalently closed circular form, is unaccounted for in this model, but serves a central role in the replication process according to this model. Form iv DNA, which is presumably derived from the minichromosome in the nucleus (7), is transcriptionally active and presumably encodes the synthesis of 35S template RNA. We do not know yet whether the supercoiled form of CaMV DNA can replicate directly, through a DNA dependent DNA synthetic process. But if it cannot, the replication of viral DNA would depend entirely on the production of the RNA template, which presumably would be destroyed during each round of DNA replication.

Hull and Covey (15) have stated that some of the CaMV DNA fragments they observe most likely arise from double-strand breaks at the single-strand break sites in the viral genome and, therefore, are probably not replication intermediates. In our experiments the only such form that might represent such a breakdown product is form iii, the double strand half-genome molecule. However, we would argue that this form more likely represents a replication intermediate, because if it were simply a broken viral DNA molecule one might expect to find both halves of the genome and only one half is found (Fig. 4). Of course, it remains to be demonstrated that these DNA molecules are true replication intermediates. It would be interesting to ask whether these forms can be isolated from viral DNA replication complexes, for example.

Reverse transcription is an important step in the replication of DNA in animal retroviruses (24) and hepatitis B virus group (25). The model of CaMV DNA replication resembles closely that of retroviruses in that first strand DNA synthesis is primed by a tRNA primer near the 5'-end of full length RNA template (24). However, there are many differences between the replication strategies of CaMV and retroviruses. First, retrovirus DNA copies integrate into the host genome. There is no evidence for CaMV DNA insertion into plant genomes, instead CaMV is found as an independent "minichromosome" in the plant

cell nucleus (7). Second, unlike retroviruses, but like hepatitis B-like virus, CaMV particles carry DNA, not RNA, and in both viruses the DNA strands are discontinuous (25). In CaMV an obligatory first step in DNA replication must be the closure of the site-specific, single-strand DNA breaks to permit transcription of the virus genome. The breaks must be closed because the break site in the d-strand interferes with the synthesis of the 35S template RNA (10,11,26). Third, retroviruses carry their own reverse transcriptase and no such activity has been found in CaMV particles. It is presumed that such an activity is encoded by the virus and must be synthesized in the infected plant cell before reverse transcription can occur.

The unusual strategy suggested for CaMV DNA replication may provide an ad hoc explanation for some curious phenomenology associated with the virus. First, CaMV genomes recombine readily (27). Among animal viruses, retroviruses show a much greater tendency to undergo recombination than DNA or RNA viruses (see 24). Second, plants can be infected with copies of the entire genome cloned in bacterial plasmids where the break sites have been sealed by cloning. In the CaMV DNA replication model proposed, the breaks are not required for, but are the products of DNA replication. Third, although "naked" viral DNA can be used to directly inoculate plants, plant protoplasts in culture are not efficiently transfected with viral DNA (a personal observation). It may be that a different form of CaMV DNA, such as the covalently closed form, is needed to initiate viral replication in protoplasts at high frequency.

ACKNOWLEDGMENTS

This work was supported by the Science and Education Administration, the U.S. Department of Agriculture under Grant 5901-04010-8-0178-0 and The National Science Foundation PCM-8141755.

*Address all correspondence to: Dr. Stephen H.Howell, c/o Laboratory of Dr. W.J.Peacock, CSIRO, Division of Plant Industry, P.O. Box 1600, Canberra City, ACT 2601, Australia

REFERENCES

1. Franck, A., Guilley, H., Jonard, G., Richards, K. and Hirth, L. (1980) Cell 21, 285-294.
2. Volovitch, M., Drugeon, G. and Yot, P. (1978) Nucleic Acids Res. 5, 2913-2925.
3. Hull, R. and Howell, S.H. (1978) Virology 86, 482-493.
4. Hull, R. (1980) Virology 100, 76-90.
5. Richards, K.E., Guilley, H. and Jonard, G. (1981) FEBS Letters 134, 67-70.
6. Hull, R., Covey, S.N., Stanley, J. and Davies, J.W. (1979) Nucleic Acids

Res. 7, 669-677.
7. Olszewski, N., Hagen, G. and Guilfoyle, T.J. (1982) Cell 29, 395-402.
8. Menissier, J., Lebeurier, G. and Hirth, L. (1982) Virology 117, 322-328.
9. Howell, S.H. and Hull, R. (1978) Virology 86, 468-481.
10. Guilley, H., Dudley, R.K., Jonard, G., Balazs, E. and Richards, K.E. (1982) Cell 30, 763-773.
11. Covey, S.N., Lomonossoff, G.P. and Hull, R. (1981) Nucleic Acids Res. 9, 6735-6747.
12. Hull, R. and Covey, S.N. (1983) Trends Biochem. Sci. 8, 119-121.
13. Guilley, H., Richards, K. and Jonard, G. (1983) EMBO J. 2, 277-282.
14. Covey, S.N., Turner, D. and Mulder, G. (1983) Nucleic Acids Res. 11, 251-264.
15. Hull, R. and Covey, S.N. (1983) Nucleic Acids Res. 11, 1881-1895.
16. Howell, S.H., Walker, L.L. and Dudley, R.K. (1980) Science 208, 1265-1267.
17. Bailey, J. and Davidson, N. (1976) Anal. Biochem. 70, 75-85.
18. Southern, E. (1979) in Methods in Enzymology 68 ed. R. Wu (Academic Press, New York) pp 152-176.
19. Wahl, G., Stern, M. and Stark, G. (1979) Proc. Natl. Acad. Sci. USA 76, 3683-3687.
20. Rigby, P.W.J., Dieckmann, M., Rhodes, C. and Berg, P. (1977) J. Mol. Biol. 113, 237-251.
21. Howell, S.H., Walden, R.M. and Marco, Y. (in press) ARCO-UCLA Winter Symposium.
22. Gilboa, E., Mitra, S.W., Goff, S. and Baltimore, D. (1979) Cell 18, 93-100.
23. Covey, S.N., Turner, D. and Mulder, G. (1983) Nucleic Acids Res. 11, 251-264.
24. Varmus, H.E. (1982) Science 216, 812-820.
25. Summers, J. and Mason, W.S. (1982) Cell 29, 403-415.
26. Dudley, R.K., Odell, J.T. and Howell, S.H. (1982) Virology 117, 19-28.
27. Walden, R.M. and Howell, S.H. (1982) Mol. and Appl. Gen. 1, 447-456.